Least Squares Orthogonal Polynomials and some applications

Claude Brezinski and Ana C. Matos
Laboratoire d'Analyse Numérique et d'Optimisation
Université des Sciences et Techniques de Lille Flandres-Artois
59655 Villeneuve d'Ascq cedex - France
e–mail:brezinsk@ano.univ–lille1.fr, matos@ano.univ–lille1.fr

Keywords: orthogonal polynomials, least squares, Padé-type approximation, quadrature methods.

1 Introduction

Let c be the linear functional on the space of complex polynomials defined by

$$\begin{aligned} c(x^i) &= c_i \in \mathbb{C}, \quad i = 0, 1, ... \\ &= 0, \qquad\quad i < 0. \end{aligned}$$

It is said that $\{P_k\}$ forms a family of (formal) orthogonal polynomials with respect to c if $\forall k$

– P_k has the exact degree k,

– $c(x^i P_k(x)) = 0$ for $i = 0, \cdots, k-1$.

Such a family exists if, $\forall k$, the Hankel determinant

$$H_k^{(0)} = \begin{vmatrix} c_0 & c_1 & \cdots & c_{k-1} \\ c_1 & c_2 & \cdots & c_k \\ \cdots & \cdots & \cdots & \cdots \\ c_{k-1} & c_k & \cdots & c_{2k-2} \end{vmatrix}$$

is different from zero. Such polynomials enjoy most of the properties of the usual orthogonal polynomials, when the functional c is given by

$$c(x^i) = \int_a^b x^i d\alpha(x),$$

where α is bounded and non decreasing in $[a, b]$ (see [1] for these properties). In this paper we study the polynomials R_k such that

$$\sum_{i=0}^{m} [c(x^i R_k(x))]^2$$

is minimized, where m is an integer strictly greater than $k-1$ (since, for $m = k-1$, we recover the previous formal orthogonal polynomials) and which can possibly depend on k. They will be called *least squares (formal) orthogonal polynomials*. They depend on the value of m but for simplicity this dependence will not be indicated in our notations .

Such polynomials arise naturally in problems of Padé approximation for power series with perturbed coefficients, and in Gaussian quadrature (as described in the last section). Some properties of these polynomials are derived, together with a recursive scheme for their computation.

2 Existence and uniqueness

Since the polynomials R_k will be defined apart from a multiplying factor, and since it is asked that the degree of R_k is exactly k we shall write

$$R_k(x) = b_0 + b_1 x + \cdots + b_{k-1}x^{k-1} + b_k x^k \quad \text{with } b_k = 1.$$

We set

$$\Phi(b_0, \cdots, b_{k-1}) = \sum_{i=0}^{m} [c(x^i R_k(x))]^2$$

and we seek for the values of $b_0, \cdots, b_{k-1}$ that minimize this quantity. That is, such that

$$\partial\Phi/\partial b_j = 0 \text{ for } j = 0, \cdots, k-1. \tag{1}$$

Setting $\gamma_n = (c_n, \cdots, c_{n+m})^T$ this system can be written

$$b_0(\gamma_0, \gamma_j) + \cdots + b_{k-1}(\gamma_{k-1}, \gamma_j) = -(\gamma_k, \gamma_j), \quad j = 0, \cdots, k-1. \tag{2}$$

Thus R_k exists and is unique if and only if the matrix A_k of this system is non singular. Setting $X = (1, x, \cdots, x^{k-1})$ and calling the right hand side of the preceding system γ we see that

$$R_k(x) = \frac{\begin{vmatrix} A_k & \gamma \\ X & x^k \end{vmatrix}}{|A_k|}.$$

If we set

$$B_k = \begin{pmatrix} c_0 & \cdots & \cdots & c_{k-1} \\ \cdots & \cdots & \cdots & \cdots \\ c_m & \cdots & \cdots & c_{m+k-1} \end{pmatrix}$$

then $A_k = B_k^T B_k$, $\gamma = B_k^T \gamma_k$ and we recover the usual solution of a system of linear equations in the least squares sense.

3 Computation

The polynomials R_k can be recursively computed by inverting the matrix A_k of the above system (2) by the bordering method, see [5]. This method is as follows. Set

$$A_{k+1} = \begin{pmatrix} A_k & u_k \\ v_k & a_k \end{pmatrix}$$

where u_k is a column vector, v_k a row vector and a_k a scalar. We then have

$$A_{k+1}^{-1} = \begin{pmatrix} A_k^{-1} + A_k^{-1} u_k \beta_k^{-1} v_k A_k^{-1} & -A_k^{-1} u_k \beta_k^{-1} \\ -\beta_k^{-1} v_k A_k^{-1} & \beta_k^{-1} \end{pmatrix}$$

where $\beta_k = a_k - v_k A_k^{-1} u_k$.

Instead of choosing the normalization $b_k = 1$ we could impose the condition $b_0 = 1$. In that case we have the system

$$b_1'(\gamma_1, \gamma_j) + \cdots + b_k'(\gamma_k, \gamma_j) = -(\gamma_0, \gamma_j), \quad j = 1, \cdots, k \tag{3}$$

and the bordering method can be used not only for computing the inverses of the matrices of the system recursively but also for obtaining its solution, since the new right hand side contains the previous one.

Let A_k' be the matrix of (3) and d_k' be the right hand side. We then have

$$A_{k+1}' = \begin{pmatrix} A_k' & u_k' \\ v_k' & a_k' \end{pmatrix} \quad d_{k+1}' = \begin{pmatrix} d_k' \\ f_k' \end{pmatrix}$$

with

$$\begin{aligned} u_k' &= ((\gamma_{k+1}, \gamma_1), \cdots, (\gamma_{k+1}, \gamma_k))^T ; \\ v_k' &= ((\gamma_1, \gamma_{k+1}), \cdots, (\gamma_k, \gamma_{k+1})) ; \\ a_k' &= (\gamma_{k+1}, \gamma_{k+1}); \\ d_k' &= ((\gamma_0, \gamma_1), \cdots, (\gamma_0, \gamma_k))^T ; \\ f_k' &= (\gamma_0, \gamma_{k+1}). \end{aligned}$$

Setting $z_k' = \left(b_1', \cdots, b_k'\right)^T$ we have

$$z_{k+1}' = \begin{pmatrix} z_k' \\ 0 \end{pmatrix} + \frac{f_k' - v_k' z_k'}{\beta_k'} \begin{pmatrix} -A_k'^{-1} u_k' \\ 1 \end{pmatrix}$$

with $\beta_k' = a_k' - v_k' A_k'^{-1} u_k'$.

Of course the bordering method can only be used if β_k (or β_k' in the second case) is different from zero. If it is not the case, instead of adding one new row and one new column to the system it is possible to add several rows and columns until a non singular β_k (which is now a square matrix) has been found (see [4] and [3]).

4 Location of the zeros

We return to the normalization $b_k = 1$. As

$$c(x^i R_k(x)) = b_0 c_i + \cdots + b_k c_{i+k} \text{ and } \frac{\partial c(x^i R_k(x))}{\partial b_j} = c_{i+j},$$

from 1 we obtain

$$\sum_{i=0}^{m} c(x^i R_k(x)) c_{i+j} = 0 \text{ for } j = 0, \cdots, k-1. \tag{4}$$

This relation can be written as

$$c(R_k(x)(c_i + c_{i+1}x + \cdots + c_{i+m}x^m)) = 0, \quad i = 0, \cdots, k-1.$$

Let us now assume that

$$c_i = \int_a^b x^i d\alpha(x), \quad i = 0, 1, \cdots$$

with α bounded and nondecreasing in $[a, b]$. We have

$$\begin{aligned} c_i + c_{i+1}x + \cdots + c_{i+m}x^m &= \sum_{j=0}^{m} \left[\int_a^b y^{i+j} d\alpha(y) \right] x^j \\ &= \int_a^b y^i \left(\sum_{j=0}^{m} x^j y^j \right) d\alpha(y). \end{aligned}$$

Set

$$w(x, \mu) = \int_a^b y^\mu \left(\sum_{j=0}^{m} x^j y^j \right) d\alpha(y).$$

Thus

$$w(x, i) = c_i + c_{i+1}x + \cdots + c_{i+m}x^m$$

and it follows that

$$c(R_k(x)w(x, i)) = \int_a^b R_k(x)w(x, i)d\alpha(x) = 0 \quad \text{for} \quad i = 0, \cdots, k-1$$

which shows that the polynomial R_k is biorthogonal in the sense of [7, 8]. Let us now study the location of the zeros of R_k. For that purpose we shall apply Theorem 3 of [7], also given as Theorem 5 of [8]. Set

$$d\Phi(x, \mu) = w(x, \mu)d\alpha(x)$$

and

$$I_k(\mu) = \int_a^b x^k d\Phi(x, \mu), \quad k = 0, 1, \cdots$$

In our case, μ takes the values $\mu_i = i - 1, \ i = 1, 2, \cdots$. Thus

$$det\left[I_i(\mu_j)\right] = det\left[(\gamma_{j-1}, \gamma_i)\right]$$

and the condition of regularity of [7, 8] is equivalent to our condition for the existence and uniqueness of R_k. According to [7, 8], we now have to look at the interpolation property of w. We have

$$w(x_i, \mu_j) = (\gamma_{j-1}, X_i)$$

where $X_i = (1, x_i, \cdots, x_i^m)^T$, the x_i's being arbitrary distinct points in $[a, b]$, and thus

$$\begin{vmatrix} (\gamma_0, X_1) & (\gamma_1, X_1) & \cdots & (\gamma_{k-1}, X_1) \\ (\gamma_0, X_2) & (\gamma_1, X_2) & \cdots & (\gamma_{k-1}, X_2) \\ \cdots & \cdots & \cdots & \cdots \\ (\gamma_0, X_k) & (\gamma_1, X_k) & \cdots & (\gamma_{k-1}, X_k) \end{vmatrix} = det(\mathcal{X}_k \Gamma_k)$$

with $\mathcal{X}_k = \begin{pmatrix} X_1^T \\ \vdots \\ X_k^T \end{pmatrix}$ and $\Gamma_k = (\gamma_0, \cdots, \gamma_{k-1})$.

The interpolation property holds if and only if $det(\mathcal{X}_k\Gamma_k) \neq 0$, that is, if and only if the matrix $\mathcal{X}_k\Gamma_k$ has rank k. Thus, using thre theorem of [7, 8], we have proved the following result

Theorem 1 *If A_k is regular and if $\mathcal{X}_k\Gamma_k$ has rank k, then R_k exists and has k distinct zeros in $[a, b]$.*

Remark: When $0 \leq a < b$, it can be proved that $det(\mathcal{X}_k\Gamma_k) \neq 0$ (see [2] for the details).

5 Applications

Our first application deals with Padé-type approximation. Let v_k be an arbitrary polynomial of degree k and let $w_k(t) = a_0 + \cdots + a_{k-1}t^{k-1}$ be defined by

$$a_i = c(x^{-i-1}v_k(x)) \quad i = 0, \cdots, k-1.$$

We set

$$\tilde{v}_k(t) = t^k v_k(t^{-1}) \text{ and } \tilde{w}_k(t) = t^{k-1}w_k(t^{-1}).$$

Let f be the formal power series

$$f(t) = \sum_{i=0}^{\infty} c_i t^i.$$

Then it can be proved that

$$f(t) - \tilde{w}_k(t)/\tilde{v}_k(t) = O(t^k) \quad (t \to 0).$$

The rational function $\tilde{w}_k(t)/\tilde{v}_k(t)$ is called a Padé-type approximant of f and it is denoted by $(k-1/k)_f(t)$, [1]. Moreover it can also be proved that

$$\begin{aligned} f(t) - \frac{\tilde{w}_k(t)}{\tilde{v}_k(t)} &= \frac{t^k}{\tilde{v}_k(t)} c\left(\frac{v_k(x)}{1-xt}\right) \\ &= \frac{t^k}{\tilde{v}_k(t)} c\left(\left(1 + xt + \cdots + x^{k-1}t^{k-1} + \frac{x^k t^k}{1-xt}\right) v_k(x)\right). \end{aligned}$$

That is

$$f(t)\tilde{v}_k(t) - \tilde{w}_k(t) = t^k \sum_{i=0}^{\infty} c(x^i v_k(x)) t^i.$$

Thus if the polynomial v_k, which is called the generating polynomial of $(k-1/k)$, satisfies

$$c(x^i v_k(x)) = 0 \text{ for } i = 0, \cdots, k-1$$

then

$$f(t) - \tilde{w}_k(t)/\tilde{v}_k(t) = O(t^{2k}).$$

In this case v_k is the formal orthogonal polynomial P_k of degree k with respect to c and $\tilde{w}_k(t)/\tilde{v}_k(t)$ is the usual Padé approximant $[k-1/k]$ of f.

As explained in [10], Padé approximants can be quite sensitive to perturbations on the coefficients c_i of the series f. Hence the idea arises to take as v_k the least squares orthogonal polynomial R_k of degree k instead of the usual orthogonal polynomial, an idea which in fact motivated our study. Of course such a choice decreases the degree of approximation, since the approximants obtained are only of the Padé-type, but it can increase the stability properties of the approximants and also their precision since $\sum_{i=0}^{m} \left[c(x^i v_k(x))\right]^2$ is minimized by the choice $v_k = R_k$. We give a numerical example that illustrates this fact.

We consider the function

$$f(z) = \frac{ln(1+z)}{z} = \sum_{i=0}^{\infty} c_i z^i$$

and we assume that we know the coefficients c_i with a certain precision. For example, we know approximate values c_i^* such that

$$| c_i - c_i^* | \leq 10^{-8}, \quad i = 0, 1, \cdots.$$

In the following table we compare the number of exact figures given by the Padé approximant with those of the least squares Padé-type approximant, both computed with the same number of coefficients c_i^*. We can see that the least squares Padé-type approximant has better stability properties.

z	Padé approximant [7/8]	LS Padé-type approximant [6/7] ($m = 8$)
1.5	6.7	7.7
1.9	5.7	7.0
2.1	5.2	6.7

Another application concerns quadrature methods. We have already shown that if the functional c is given by

$$c_i = \int_a^b x^i d\alpha(x), \quad i = 0, 1, \cdots \quad 0 \leq a < b$$

with α bounded and nondecreasing, then the corresponding least squares orthogonal polynomial of degree k, R_k, has k distinct zeros in $[a, b]$. We can then construct quadrature formulae of the interpolatory type.

If $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the zeros of R_k, we can approximate the integral

$$I = \int_a^b f(x) d\alpha(x)$$

by

$$I_k = A_1 f(\lambda_1) + A_2 f(\lambda_2) + \cdots + A_k f(\lambda_k) \tag{5}$$

where

$$A_i = \int_a^b \frac{\pi(x)}{\pi'(\lambda_i)(x - \lambda_i)} d\alpha(x), \quad \pi(x) = \prod_{j=1}^{k} (x - \lambda_j).$$

This corresponds to replacing the function f by its interpolating polynomial at the knots $\lambda_1, \cdots, \lambda_k$. The truncation error of (5) is given by

$$I - I_k = E_T = \int_a^b f[\lambda_1, \cdots, \lambda_k, x] R_k(x) d\alpha(x).$$

Expanding the divided difference we see

$$\begin{aligned} & f[\lambda_1, \cdots, \lambda_k, x] = \\ = & \sum_{i=1}^{k} f[\lambda_1, \cdots, \lambda_k, \lambda_{k+1}, \cdots, \lambda_{k+i}](x - \lambda_{k+1}) \cdots (x - \lambda_{k+i-1}) \\ + & f[\lambda_1, \cdots, \lambda_k, \lambda_{k+1}, \cdots, \lambda_{k+m+1}, x](x - \lambda_{k+1}) \cdots (x - \lambda_{k+m+1}) \end{aligned}$$

for $\lambda_{k+1}, \cdots, \lambda_{k+m+1}$ any points in the domain of definition $\mathcal{D}_f$ of f. If $0 \in \mathcal{D}_f$, then we can choose

$$\lambda_{k+1} = \lambda_{k+2} = \cdots = \lambda_{k+m+1} = 0.$$

Setting

$$M_i = f[\lambda_1, \cdots, \lambda_{k+i}]$$

we get

$$f[\lambda_1, \cdots, \lambda_k, x] = \sum_{i=1}^{m+1} M_i x^{i-1} + x^{m+1} f[\lambda_1, \cdots, \lambda_{k+m+1}, x]$$

and hence, for the truncation error

$$E_T = \sum_{i=0}^{m} M_{i+1} \left(\int_a^b R_k(x) x^i d\alpha(x) \right) + \int_a^b f[\lambda_1, \cdots \lambda_{k+m+1}, x] x^{m+1} R_k(x) d\alpha(x)$$

$$\text{with } \sum_{i=0}^{m} \left(\int_a^b R_k(x) x^i d\alpha(x) \right)^2 \quad \text{minimised.}$$

Moreover, if $f \in C^{k+m+1}([a,b])$ and, since x^{m+1} is positive over $[a,b]$, we obtain

$$\int_a^b f[\lambda_1, \cdots, \lambda_{k+m+1}, x] x^{m+1} R_k(x) d\alpha(x) =$$
$$= \frac{c_{m+1}}{(k+m+1)!} R_k(\lambda) f^{(k+m+1)}(\eta), \quad \lambda, \eta \in [a,b]$$

and, for the error,

$$E_T = \sum_{i=0}^{m} \frac{f^{(k+i)}(\eta_i)}{(k+i)!} \left(\int_a^b R_k(x) x^i d\alpha(x) \right) + \frac{c_{m+1}}{(k+m+1)!} R_k(\lambda) f^{(k+m+1)}(\eta) \tag{6}$$

$$\eta_i \in [a,b] \quad i = 0, \cdots, m; \quad \lambda, \eta \in [a,b].$$

We remark that in the case where $m = k-1$, R_k is the orthogonal polynomial with respect to the functional c and so formula (5) corresponds to a Gaussian quadrature formula. An advantage of the quadrature formulae (5) is that they are less sensitive to perturbations on the sequence of moments c_i, as is shown in the following numerical example. Such a case can arise in some applications where the formula giving the moments c_i is sensitive to rounding errors, see [11] for example.

Consider the functional c defined by

$$c_i = \int_0^1 x^i dx = \frac{1}{i+1}$$

and perturb the coefficients in the following way

i	c_i^*	i	c_i^*
0	1.00000011	6	0.14285700
1	0.50000029	7	0.12500000
2	0.33333340	8	0.11111109
3	0.25000101	9	0.10000000
4	0.20000070	10	0.09090899
5	0.16666600	11	0.08333300

We can construct from these coefficients the least squares orthogonal polynomials and the corresponding quadrature formulae (5). The precision of the numerical approximations of $I = \int_0^1 f(x)dx$ is given in the following table

$f(x)$	$k=5;\ m=4$ Gaussian quad	$k=5;\ m=6$ least squares quad
$1/(x+0.5)$	$-2.2*10^{-5}$	$-6.2*10^{-6}$
$1/(x+0.3)$	$-2.1*10^{-4}$	$-1.2*10^{-5}$

We can obtain other applications from the following generalization. Instead of minimizing $\sum_{i=0}^{m}\left[c(x^i R_k(x))\right]^2$ we can introduce weights and minimize

$$\Phi^*(b_0,\cdots,b_{k-1}) = \sum_{i=0}^{m} p_i \left[c(x^i R_k^*(x))\right]^2$$

with $p_i > 0,\ i = 0,\cdots,m$. If we choose the inner product

$$(\gamma_i,\gamma_j)^* = \sum_{k=0}^{m} p_k c_{i+k} c_{j+k}$$

the solution of this problem can be computed as in the previous case and all the properties of the polynomials are still true. It can be seen, from numerical examples, that if the sequence of moments c_i has a decreasing precision, we can expect that the least squares Padé-type approximants constructed with a decreasing sequence of weights will give a better result. In the same way, for the quadrature formulae (5), from the expression (6) of the truncation error and the knowledge of the magnitude of the derivatives, we can reduce this error by choosing appropriate weights. Some other possible applications of least squares orthogonal polynomials will be studied in the future.

References

[1] C.Brezinski, *Padé Type Approximation and General Orthogonal Polynomials*, ISNM Vol. 50, Birkhäuser Verlag, Basel, 1980.

[2] C. Brezinski, Ana C. Matos, *Least Squares Orthogonal Polynomials*, J. Comput. Appl. Math., 46 (1993) 229–239.

[3] C.Brezinski, M. Redivo Zaglia, H.Sadok, *A breakdown-free Lanczos type algorithm for solving linear systems* , to appear.

[4] C.Brezinski, M.Redivo Zaglia. *Extrapolation Methods. Theory and Practice*, North-Holland, Amsterdam, 1991.

[5] V.N.Faddeeva, *Computational Methods of Linear Algebra*, Dover, New York, 1959.

[6] F.R.Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1959.

[7] A.Iserles, S.P.Nørsett, *Bi-orthogonal polynomials*, in *Orthogonal Polynomials and Their Applications*, C.Brezinski, A.Draux, A.P.Magnus, P.Maroni and A.Ronveaux eds., LNM 1171, Springer-Verlag, Berlin, 1985, pp.92–100.

[8] A.Iserles, S.P.Nørsett, *On the theory of biorthogonal polynomials*, Trans. Amer. Math. Soc., 306 (1988) 455–474.

[9] S.Karlin, *Total Positivity*, Stanford University Press, Stanford, 1968.

[10] J.C.Mason, *Some applications and drawbacks of Padé Approximants*, in *Approximation Theory and Applications*, Zvi Ziegler ed., Academic Press, New York, 1981, pp.207–223.

[11] M.Morandi Cecchi, M. Redivo Zaglia, *A new recursive algorithm for a Gaussian quadrature formula via orthogonal polynomials* ,in *Orthogonal polynomials and their Applications*, C. Brezinski, L. Gori and A. Ronveaux (editors), J.C. Baltzer AG, IMACS, 1991, pp.353–358.